Semantic-Web Technology: Applications at NASA

Naveen Ashish
Research Institute for Advanced Computer Science
NASA Ames Research Center, MS 269/3, Moffett Field CA94035
USA
ashish@email.arc.nasa.gov
http://ic.arc.nasa.gov

Abstract. We provide a description of work at the National Aeronautics and Space Administration (NASA) on building systems based on semantic-web concepts and technologies. NASA has been one of the early adopters of semantic-web technologies for practical applications. Indeed there are several ongoing (IT) endeavors on building semantics based systems for use in diverse NASA domains ranging from collaborative scientific activity to accident and mishap investigation to enterprise search to scientific information gathering and integration to aviation safety decision support. We provide a brief overview of many applications and ongoing work with the goal of informing the external community of these NASA endeavors.

1 Introduction

We present an overview of several completed and ongoing endeavors at NASA of systems based on semantic web technologies. NASA[1] as an agency has a strategic focus on the four areas of Exploration, Space, Science and Aeronautics. Information Technology(IT) R&D at NASA focuses on the areas on intelligent systems, supercomputing and networking to meet NASA mission needs. Besides, NASA also conducts R&D activities for aviation information management and decision support in collaboration with the Federal Aviation Administration (FAA). The focus of Semantic Web related work at NASA is not so much on 'fundamental' Semantic Web issues such as developing languages for ontologies or markup, or developing tools such as for ontology matching, rather the focus is heavily towards applications based on these technologies that can benefit NASA missions. We describe a variety of systems including systems for collaborative knowledge sharing, taxonomy and enterprise search, scientific information gathering, scientific discovery, and aviation information management that have been or are being built to address various NASA (IT) challenges.

[1] http://www.nasa.gov

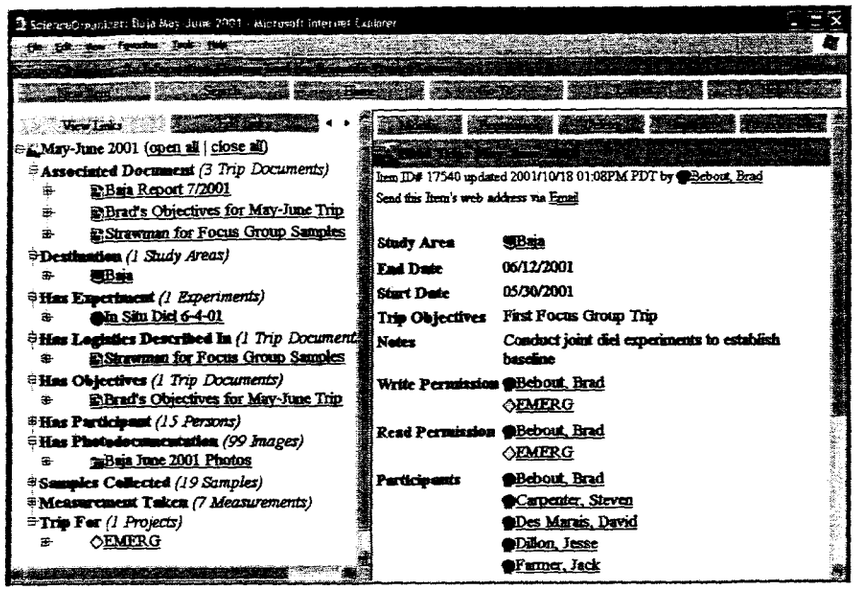

Fig 1. SemanticOrganizer

- Principled ontology evolution (over the course of an application) is difficult to sustain
- Navigating a large semantic network (5000 nodes, 30000-50000 semantic connections) is problematic for end users
- Automated metadata extraction is critical as users or administrators cannot manually annotate or enrich the data coming in so that it may be semantically related and queried.

More information about the SemanticOrganizer may be obtained at http://ic.arc.nasa.gov/sciencedesk/. The points-of-contact(POCs) for the project are Dr. Richard Keller (keller@email.arc.nasa.gov) and Dr. Daniel Berrios (berrios@email.arc.nasa.gov) both at NASA Ames Research Center.

2.2 NASA Taxonomy

The NASA Taxonomy is an effort on building an enterprise-wide taxonomy for NASA. The intended use of the taxonomy is to help NASA personnel – scientists and engineers find information, through the use of intelligent search, browsing, and navigation systems that utilize the taxonomy. The NASA Taxonomy development, funded out of the NASA CIO office, was led by Taxonomy Strategies Inc.[2] A top-down approach to taxonomy development was followed. Subject matter experts from various areas such as unmanned space mission development, mission technology development, engineering configuration management and product data management systems were extensively consulted. Also, input was sought from managers of IT systems and project content for manned missions. The taxonomy development was done in accordance with industry best practices such as hierarchical granularity,

[2] http://www.taxonomystrategies.com

DAML+OIL and are now being updated to OWL. They are about various domains and areas such as ontologies about the earth realm (the various "spheres" i.e., atmosphere, ocean, solid earth etc. of the earth realm), substances (particles, chemical compounds etc.), living elements (plants and animal species), physical properties, unites, numerical entities, temporal entities and relationships, spatial entities and relationships, natural phenomena and human activities. The Postgres ORDBMS has been used for ontology storage given its support for object-oriented features as well as geospatial query processing capabilities. A search tool that utilizes these ontologies and also using Latent Semantic Analysis (LSA) to uncover relationships between (search) terms is being developed.

Information about SWEET is available at http://sweet.jpl.nasa.gov and the POC is Dr. Robert Raskin (raskin@seastar.jpl.nasa.gov) at JPL.

2.4 NASA Discovery Systems Project

In addition to completed or ongoing R&D projects, we also describe an envisioned new NASA research program which is strongly related to semantic interoperation amongst other things. The NASA Discovery Systems initiative is a planned 5 year effort that seeks to revolutionize scientific activity. The scientific discovery process at NASA (and indeed in general) is highly data driven. Scientists and engineers have a significant need to understand the vast data sources that are being created through various NASA technology and projects. The current process to integrate and analyze data is labor intensive and requires expert knowledge about data formats and archives. Current discovery and analysis tools are fragmented and mainly support a single person working on small, clean data sets in restricted domains. The Discovery Systems project will develop and demonstrate technologies to handle the details and provide ubiquitous and seamless access to and integration of increasingly massive and diverse information from distributed sources. New technology that generates explanatory, exploratory, and predictive models, makes these tools easier to use, and integrates them in interactive, exploratory environments that let scientists and engineers formulate and solve increasingly complex interdisciplinary problems is sought. Clearly the importance of semantic technologies cannot be underscored
The following are the identified thrust areas that have been identified for pursuit in realizing the Discovery Systems vision:

- Collaborative exploratory environments and knowledge sharingMachine assisted model discovery and refinementMachine integration of data based on content Distributed data search, access and analysis More information about the Discovery Systems initiative can be obtained at http://postdoc.arc.nasa.gov/ds-planning/public. The POC for Discovery Systems is Dr. Barney Pell (pell@email.arc.nasa.gov) at NASA Ames Research Center.**2.5 SWIM**

System Wide Information Management (SWIM) is an initiative[3] aimed at providing information integration capabilities for systems in the National Airspace System (NAS). The NAS is a complex network of distributed and interconnected information

2. Raskin, R. *Semantic Web for Earth and Environmental Terminology (SWEET).* in *Workshop on Semantic Web Technologies for Searching and Retrieving Scientific Data at ISWC 2003.* 2003. Sanibel Island, FL: JPL.
3. Ashish, N. and A. Goforth. *Intelligent Information Fusion in the Aviation Domain: A Semantic-Web Based Approach.* in *AIAA Intelligent Systems Technical Conference.* 2004. Chicago, IL.